

Molecular Evolution of Vertebrate Neurotrophins: Co-Option of the Highly Conserved Nerve Growth Factor Gene into the Advanced Snake Venom Arsenal

Kartik Sunagar^{1,2}, Bryan Grieg Fry^{3,4}, Timothy N. W. Jackson^{3,4}, Nicholas R. Casewell^{5,6}, Eivind A. B. Undheim^{3,4}, Nicolas Vidal⁷, Syed A. Ali^{3,8}, Glenn F. King⁴, Karthikeyan Vasudevan⁹, Vitor Vasconcelos^{1,2}, Agostinho Antunes^{1,2*}

1 CIMAR/CIIMAR, Centro Interdisciplinar de Investigação Marinha e Ambiental, Universidade do Porto, Porto, Portugal, **2** Departamento de Biologia, Faculdade de Ciências, Universidade do Porto, Porto, Portugal, **3** Venom Evolution Lab, School of Biological Sciences, The University of Queensland, St. Lucia, Queensland, Australia, **4** Institute for Molecular Bioscience, University of Queensland, St Lucia, Queensland, Australia, **5** Molecular Ecology and Evolution Group, School of Biological Sciences, Bangor University, Bangor, United Kingdom, **6** Alistair Reid Venom Research Unit, Liverpool School of Tropical Medicine, Liverpool, United Kingdom, **7** Département Systématique et Evolution, Service de Systématique Moléculaire, UMR 7138, Muséum National d'Histoire Naturelle, Paris, France, **8** HEJ Research Institute of Chemistry, International Center for Chemical and Biological Sciences (ICCBS), University of Karachi, Karachi, Pakistan, **9** Wildlife Institute of India, Dehradun, Uttarakhand, India

Abstract

Neurotrophins are a diverse class of structurally related proteins, essential for neuronal development, survival, plasticity and regeneration. They are characterized by major family members, such as the nerve growth factors (NGF), brain-derived neurotrophic factors (BDNF) and neurotrophin-3 (NT-3), which have been demonstrated here to lack coding sequence variations and follow the regime of negative selection, highlighting their extremely important conserved role in vertebrate homeostasis. However, in stark contrast, venom NGF secreted as part of the chemical arsenal of the venomous advanced snake family Elapidae (and to a lesser extent Viperidae) have characteristics consistent with the typical accelerated molecular evolution of venom components. This includes a rapid rate of diversification under the significant influence of positive-selection, with the majority of positively-selected sites found in the secreted β -polypeptide chain (74%) and on the molecular surface of the protein (92%), while the core structural and functional residues remain highly constrained. Such focal mutagenesis generates active residues on the toxin molecular surface, which are capable of interacting with novel biological targets in prey to induce a myriad of pharmacological effects. We propose that caenophidian NGFs could participate in prey-venomoming by causing a massive release of chemical mediators from mast cells to mount inflammatory reactions and increase vascular permeability, thereby aiding the spread of other toxins and/or by acting as proapoptotic factors. Despite their presence in reptilian venom having been known for over 60 years, this is the first evidence that venom-secreted NGF follows the molecular evolutionary pattern of other venom components, and thus likely participates in prey-venomoming.

Citation: Sunagar K, Fry BG, Jackson TNW, Casewell NR, Undheim EAB, et al. (2013) Molecular Evolution of Vertebrate Neurotrophins: Co-Option of the Highly Conserved Nerve Growth Factor Gene into the Advanced Snake Venom Arsenal. PLoS ONE 8(11): e81827. doi:10.1371/journal.pone.0081827

Editor: Christine A. Kozak, National Institute of Allergy and Infectious Diseases, United States of America

Received: July 8, 2013; **Accepted:** October 17, 2013; **Published:** November 29, 2013

Copyright: © 2013 Kartik Sunagar. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: KS was funded by PhD grant (SFRH/BD/61959/2009) from F.C.T (Fundação para a Ciência e a Tecnologia). BGF was funded by an Australian Research Council Future Fellowship and by the University of Queensland. TNWJ was funded by an Australian Postgraduate Award. NRC was funded by fellowship award (NE/J018678/1) from the Natural Environment Research Council, UK. EABU acknowledges funding from the University of Queensland (International Postgraduate Research Scholarship, UQ Centennial Scholarship, and UQ Advantage Top-Up Scholarship) and the Norwegian State Education Loans Fund. NV was funded by the 'Service de Systématique Moléculaire' of the Muséum National d'Histoire Naturelle [UMS 2700 (OMSI)]. GFK was funded by Australian Research Council Discovery Grant DP130103813. SAA was the recipient of postdoctoral fellowship (PDRF Phase II Batch-V) from Higher Education Commission (HEC Islamabad) Pakistan. AA was partially supported by the European Regional Development Fund (ERDF) through the COMPETE - Operational Competitiveness Programme and national funds through F.C.T under the projects PEst-C/MAR/LA0015/2013, PTDC/AAC-AMB/104983/2008 (FCOMP-01-0124-FEDER-008610) and PTDC/AAC-AMB/121301/2010 (FCOMP-01-0124-FEDER-019490). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors have declared that no competing interests exist.

* E-mail: aantunes@ciimar.up.pt

Introduction

Venom, a complex biochemical cocktail of biologically active components, such as proteins, peptides, amino acids, neurotransmitters and polyamines, has underpinned the diversification and evolutionary success of several animal lineages [1]. This key evolutionary innovation is employed by a plethora of animals for predation, competitor deterrence and defence [2–5]. The scientific consensus is that venom components originate via toxin recruitment events, as part of which physiological protein-encoding genes are duplicated and the new copies are selectively expressed in the venom gland [5–15]. Over the years, our understanding of the origin and diversification of snake venoms has greatly increased, largely due to advances in transcriptomics and proteomics [16–22]. However, the precise role of certain proteins, which are secreted as part of the biochemical venom arsenal, still remains to be elucidated. Nerve growth factor (NGF), a key member of the neurotrophin family, is one such class of protein whose presence in snake venoms has been intriguing. Since its discovery in the late 1950s, NGF has been reported from the venoms of various caenophidian (advanced) snakes, including members of the front-fanged elapid and viperid families as well as from venomous lizards [23–29], but its function and relative importance in snake venoms remains unknown [8].

Neurotrophins represent a family of structurally related proteins, crucial for neuronal development, survival, death, regeneration and plasticity. According to the classical neurotrophic hypothesis, neurotrophins are produced in limiting amounts and the survival of the innervating neurons is dependent on winning the competition for sufficient quantities of these factors [30,31]. Neurotrophins contain gene family members such as nerve growth factor (NGF), brain-derived neurotrophic factor (BDNF), neurotrophin-3 (NT-3) and neurotrophin-4/5 (NT-4/5) [32–35], all of which function by interacting with the p75 neurotrophin receptor (NTR) in their proneurotrophin forms and various structurally related tropomyosin-related kinase or Tyrosine kinase receptors (Trk) in their active cleaved form [36,37]. While NGF specifically activates TrkA [38], BDNF and NT-4/5 interact with TrkB [36,39]. NT-3 primarily interacts with TrkC and is unique in also being capable of weakly binding to both TrkA and TrkB [36].

Neurotrophins have been extensively studied not only because they are perceived as one of the primary factors responsible for the complexity of vertebrate nervous systems, but also because of their involvement in cognition and memory. Knockout of the genes encoding NGF, BDNF and NT-3 genes is fatal in mice, highlighting the importance of these proteins for survival and normal neuronal development. Not surprisingly, abnormalities associated with the production of neurotrophins have been linked with neuropathies and neurodegenerative disorders.

In order to investigate the role of NGF in the venom of *Toxicofera* reptiles [24], we have investigated the molecular evolution of these proteins in reptilian (turtles; squamates: *Laterata*, *Scinciformata*, *Gekkota*; *Toxicofera* lizards: *Anguimorpha* and *Iguania*; *Henophidia* snakes; advanced

snakes: *Elapidae*, *Viperidae* and ‘non-front-fanged’ advanced snakes) and mammalian lineages, by employing sophisticated protein and codon-level selection assessments. We further compare the molecular evolution of NGF with the other major members of the neurotrophin family, namely BDNF and NT-3, in a wide array of reptilian and mammalian lineages. Molecular evolution analyses conducted on a dataset of 1183 nucleotide sequences revealed that these genes have remained largely unchanged since their origin over 300 million years (the split between mammals and reptiles: www.timetree.org) due to the extremely important functions they play in vertebrate homeostasis.

Results

Bayesian and maximum-likelihood analyses of NGF, BDNF and NT-3 genes retrieved trees with the same topology (Figures 1–3; Figure S1–S4), which were in concordance with the earlier reported phylogenies of neurotrophins [40].

One ratio model (ORM), the simplest of the lineage-specific models, computed ω of less than 0.50, therefore indicating an influence of negative selection on NGF, BDNF and NT-3 genes in all of the mammalian and reptilian lineages examined (Tables S1–S3). This highly conservative model can only detect positive selection when the ω ratio averaged over all the sites along the lineages in a phylogenetic tree is significantly greater than one. Nevertheless, the computed ω of 0.86 for the *Elapidae* NGF highlights a relatively greater accumulation of variation in this lineage (Table S1).

To detect episodic diversifying selection, which only affects certain sites in the protein, we employed site-specific models (Tables 1–3; Tables S1–S3). Like ORM, site model 8 (M8) also indicated a strong influence of negative selection on various reptilian and mammalian NGF, BDNF and NT-3 genes (Tables 1–3; Tables S1–S3; Figures 4 and 5). However, *Elapidae* NGF was found to be rapidly evolving ($\omega = 0.95$). The Bayes Empirical Bayes (BEB) approach in M8 identified as many as 19 positively-selected sites (13% of the total sites with ω of 4.09) in *Elapidae* NGF, indicating the strong influence of positive selection on this lineage (Table 1 and Table S1). Although, M8 failed to detect variation in the *Viperidae* ($\omega = 0.77$) and ‘non-front-fanged’ advanced snake ($\omega = 0.58$) NGF lineages, as many as 9% ($\omega = 2.40$) and 8% ($\omega = 2.50$) of the total codon sites were detected as rapidly diversifying in the respective lineages. Single Likelihood Ancestral Counting (SLAC), Fixed Effects Likelihood (FEL), Random Effects Likelihood (REL), Mixed Effects Model of Evolution (MEME), Fast, Unconstrained Bayesian Approximation (FUBAR) and integrative approach, conclusively supported these findings and highlighted the complete lack of variations in NGF, BDNF and NT-3 genes in various reptilian and mammalian lineages examined (Tables 1–3). However, elapid venom NGF was found to be rapidly evolving under the influence of positive selection (integrative analyses: 34 positively selected sites). *Viperidae* NGF was also found to accumulate relatively greater variations (integrative analyses: 8 positively selected sites) in comparison to the non-venomous reptilian and mammalian NGF lineages. Analysis of NGF from the venom-glands of ‘non-

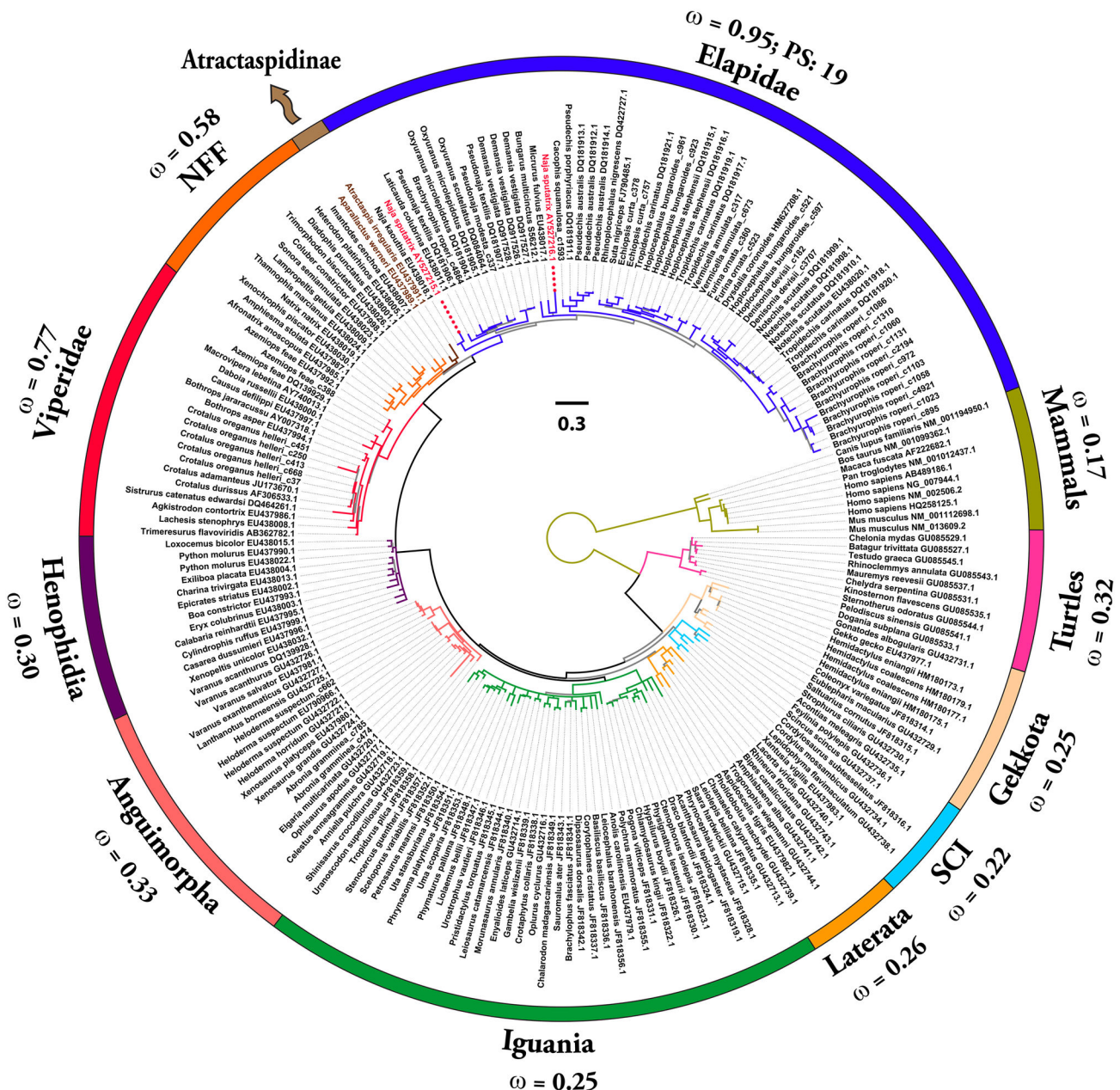


Figure 1. Bayesian molecular phylogeny of nerve growth factors (NGF). Branches with the Bayesian posterior probability (B.P.P) of less than 0.85 are highlighted in grey (remaining in colours). Site model 8 (M8) computed ω values for respective lineages are presented. The number of positively selected sites (PP \geq 0.95) detected by M8's Bayes-Empirical Bayes (BEB) approach in Elapidae lineage is also indicated. Elapid sequences representing putative duplicate genes are indicated with red labels [NFF: 'non-front-fanged' advanced snakes; SCI: Scinciformata].

doi: 10.1371/journal.pone.0081827.g001

front-fanged' advanced snakes and venomous lizards (Anguimorpha and Iguania) was hampered by the scarcity of sequences. Despite this, the integrative approach was able to identify five positively selected sites in the 'non-front-fanged' advanced snake lineages.

We further evaluated selection pressures along the elapid and viperid NGF lineages using two-ratio model (which assesses selection pressures only across lineages) and the branch-site test A (which assesses selection pressures across the sites and along the lineages). The two-ratio model failed ($p > 0.05$) to detect positive selection in elapid ($\omega = 0.85$) and

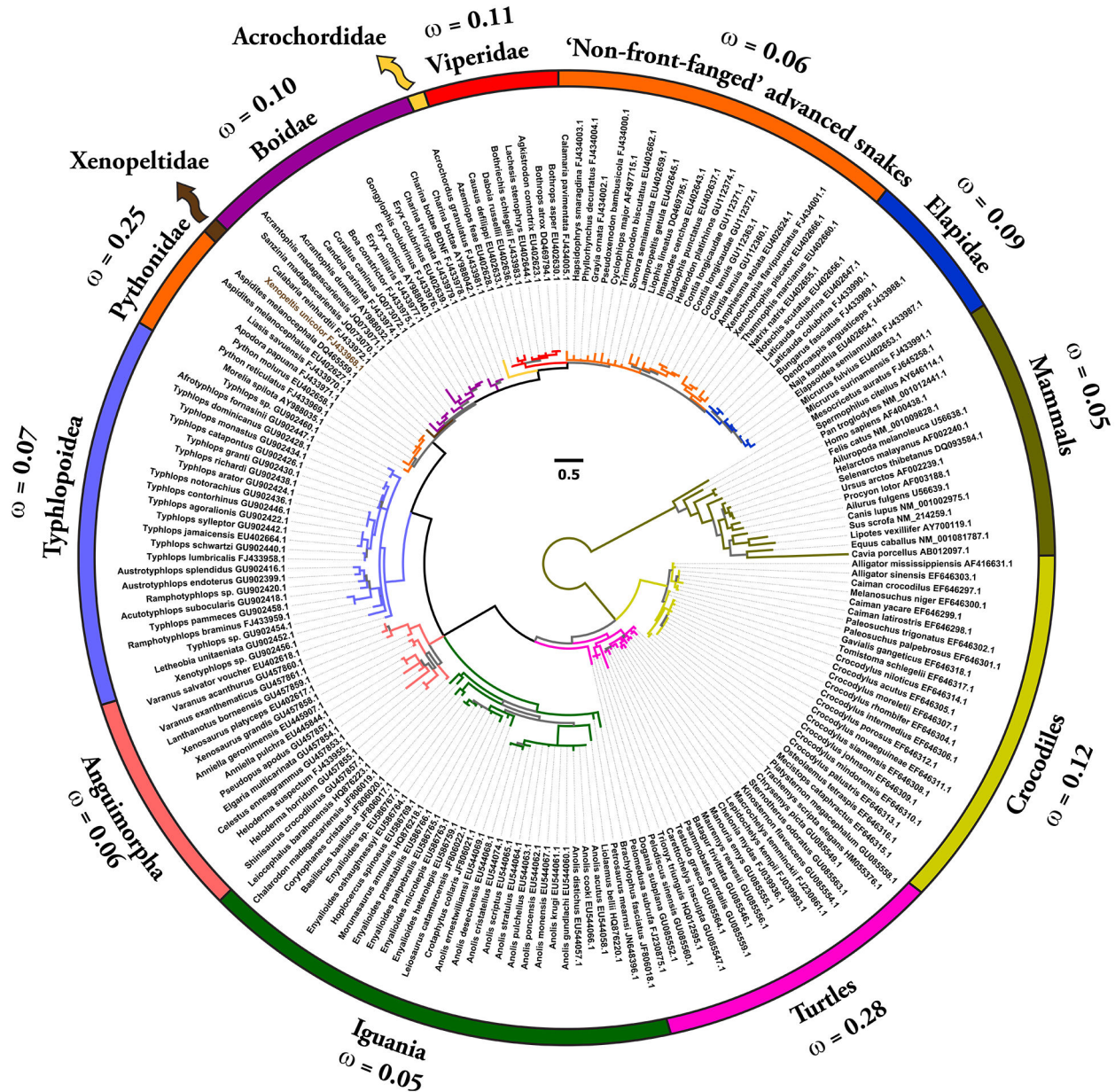


Figure 2. Bayesian molecular phylogeny of brain-derived neurotrophic factors (BDNF). Branches with the Bayesian posterior probability (B.P.P.) of less than 0.85 are highlighted in grey (remaining in colours). Site model 8 (M8) computed ω values for respective lineages are presented.

doi: 10.1371/journal.pone.0081827.g002

viperid NGF ($\omega=0.82$). However, branch-site test A indicated a strong influence of positive selection on elapid NGF [$\omega = 3.90$; 4.5% positively selected (PS) sites; significant at 0.001 after Bonferroni correction] and to a lesser extent on Viperidae NGF ($\omega = 2.01$ and 8.1% of PS-sites; significant at 0.001 after Bonferroni correction: Table S4). We further employed the clade model c approach to simultaneously compute and compare the ω values of various Toxicofera NGF lineages (Table 1). Clade model analyses indicated the influence of

positive selection in shaping the evolution of Elapidae NGF ($\omega=2.38$), while the remaining lineages were found to be constrained by negative selection (but again with the caveat that fewer 'non-front-fanged' advanced snake and Toxicofera lizard venom-gland specific NGF sequences were available; Table 1).

To derive further support for the sites detected as positively selected by the nucleotide analyses of elapid NGF, we employed a complementary amino acid-level approach

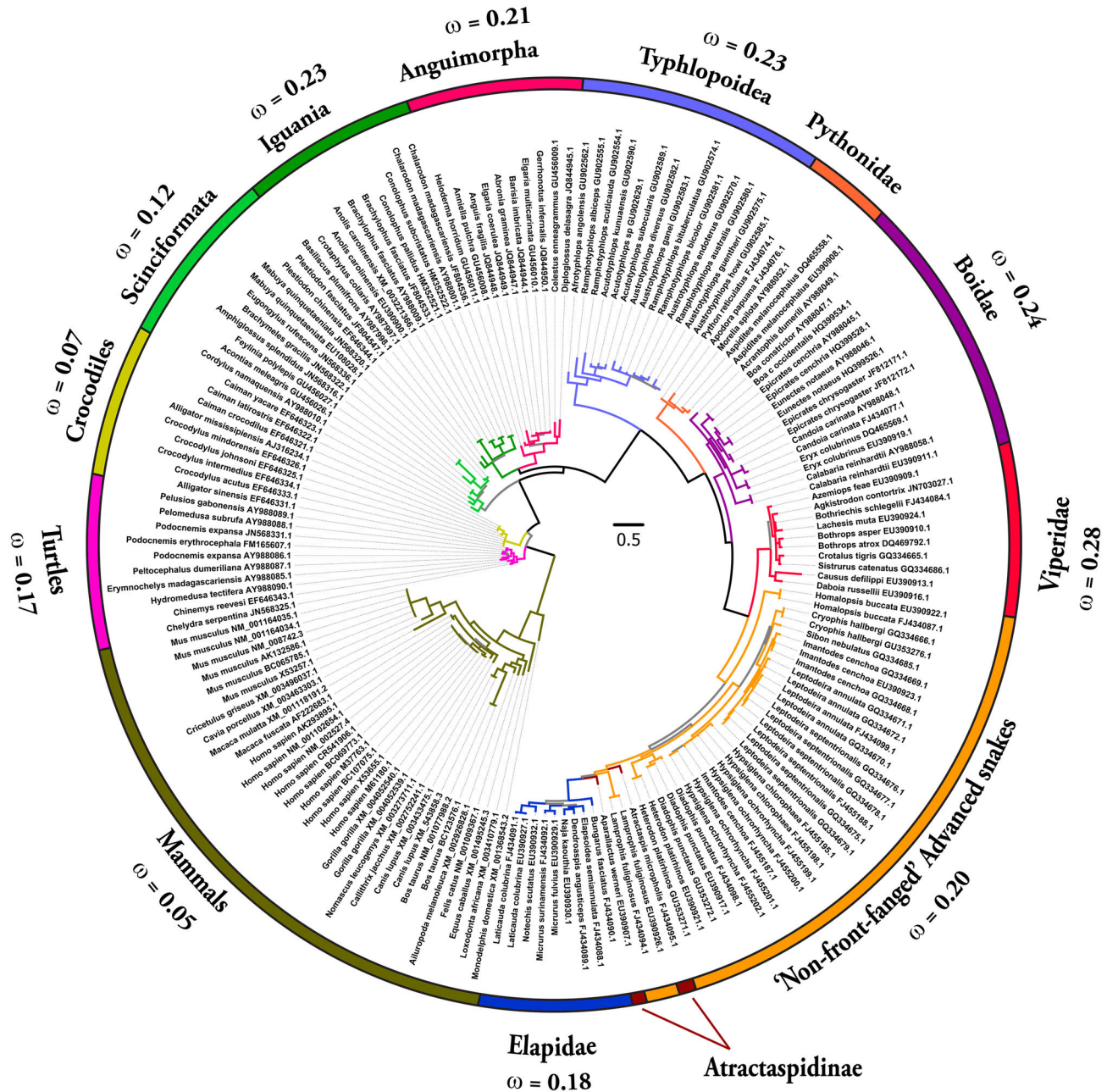


Figure 3. Bayesian molecular phylogeny of neurotrophin-3 (NT-3). Branches with the Bayesian posterior probability (B.P.P.) of less than 0.85 are highlighted in grey (remaining in colours). Site model 8 (M8) computed ω values for respective lineages are presented.

doi: 10.1371/journal.pone.0081827.g003

implemented in TreeSAAP (Table 4). Since a mutated amino acid can share similar/identical biochemical and structural properties with the ancestral amino acid being replaced, not all non-synonymous mutations may affect the structure and function of a protein. Hence, the evaluation of the influence of selection on the biochemical and structural properties of proteins provides further insights about the strength of a

mutation in affecting the fitness of an organism. With this combined approach, we were able to identify 13 positively selected sites (68% of positively selected sites) in elapid NGF (Table 4). Mapping of these mutations onto the three-dimensional structure of NGF revealed that a majority of detected hypermutable sites (74%) were concentrated in the secreted region (β -polypeptide chain) of Elapidae NGF (Figure

Table 1. Molecular evolution of Nerve Growth Factor (NGF).

	SLAC ^a	FEL ^b	REL ^c	FUBAR ^d	Integrative	MEME ^e	M8	M2a	Clade
Elapidae									
$\omega > 1^f$	2	13	28	22	34		19	15	
$\omega < 1^g$	12	20	11	18	22	21	(14+5)	(13+2)	2.38
$\omega =$	0.90	-	1.93	-	-		0.95	0.94	
Viperidae									
$\omega > 1^f$	0	0	2	0	8		0	0	
$\omega < 1^g$	3	13	0	7	13	6	-	-	1.03
$\omega =$	0.65	-	0.88	-	-		0.77	0.77	
NFF									
$\omega > 1^f$	0	1	3	2	5		0	0	
$\omega < 1^g$	1	2	0	3	3	2	-	-	0.57
$\omega =$	0.59	-	0.67	-	-		0.58	0.58	
Henophidia									
$\omega > 1^f$	0	0	0	0	1		0	0	
$\omega < 1^g$	3	12	1	13	13	1	-	-	0.0001
$\omega =$	0.30	-	0.43	-	-		0.30	0.30	
Iguania									
$\omega > 1^f$	0	0	0	0	4		0	0	
$\omega < 1^g$	35	55	73	83	87	4	-	-	0.20
$\omega =$	0.26	-	0.76	-	-		0.25	0.29	
Anguimorpha									
$\omega > 1^f$	0	0	0	0	1		0	0	
$\omega < 1^g$	4	13	0	15	15	1	-	-	0.20
$\omega =$	0.33	-	0.48	-	-		0.33	0.36	
Gekkota									
$\omega > 1^f$	0	0	3	1	5		0	0	
$\omega < 1^g$	1	15	0	16	16	1	-	-	-
$\omega =$	0.24	-	0.34	-	-		0.25	0.25	
Scinciformata									
$\omega > 1^f$	0	0	0	0	1		0	0	
$\omega < 1^g$	8	18	All	18	19	1	-	-	-
$\omega =$	0.22	-	0.26	-	-		0.22	0.25	
Laterata									
$\omega > 1^f$	0	0	0	0	0		0	0	
$\omega < 1^g$	5	12	0	14	15	0	-	-	-
$\omega =$	0.26	-	0.36	-	-		0.26	0.27	
Turtles									
$\omega > 1^f$	0	0	0	1	1		0	0	
$\omega < 1^g$	9	13	0	14	14	0	-	-	-
$\omega =$	0.33	-	0.38	-	-		0.32	0.34	
Mammals									
$\omega > 1^f$	0	1	0	0	4		0	0	
$\omega < 1^g$	24	54	All	86	86	3	-	-	-
$\omega =$	0.21	-	0.22	-	-		0.17	0.11	

S5). Mutation mapping also revealed that a greater proportion of amino acid residues (92%; remaining sites couldn't be assigned to exposed/buried class) were exposed to the surrounding medium, indicating that these Elapid venom proteins evolve through focal mutagenesis (Table 4).

Evolutionary fingerprint analyses, which is based on the probability distribution of site-to-site synonymous (α or dS) and nonsynonymous (β or dN) substitution rates in an alignment,

identified a significantly large proportion of sites in elapid NGF as evolving under the influence of positive selection, while the remaining reptilian and mammalian NGF, BDNF and NT-3 lineages were depicted as constrained by extreme negative selection pressures (Figure S6-S8). It is noteworthy that a small fraction of sites in all snake lineages (Caenophidia and Henophidia) seem to have accumulated a greater proportion of non-synonymous to synonymous mutations, although these

Table 1 (continued).**a:** Single Likelihood Ancestor Counting**b:** Fixed-effects likelihood**c:** Random-effects likelihood**d:** Fast, Unconstrained Bayesian Approximation**Integrative:** Sites detected in common by SLAC, FEL, REL, FUBAR and MEME**e:** Sites detected as experiencing episodic diversifying selection (0.05 significance) by the Mixed Effects Model Evolution (MEME) **M8:** Positively-selected sites detected using the Bayes Empirical Bayes approach implemented in M8. Sites detected at 0.99 and 0.95 significance are indicated in the parenthesis **M2a:** Positively-selected sites detected using the Bayes Empirical Bayes approach implemented in M2a. Sites detected at 0.99 and 0.95 significance are indicated in the parenthesis **Clade:** Omega computed by the clade model**f:** Number of positively selected sites at 0.05 significance (for SLAC, FEL) or 50 Bayes factor (for REL) / number of sites under pervasive diversifying selection at the posterior probability ≥ 0.9 (FUBAR)**g:** Number of negatively selected sites at 0.05 significance (for SLAC, FEL) or 50 Bayes factor (for REL) / number of sites under pervasive purifying selection at the posterior probability ≥ 0.9 (FUBAR) **ω :** mean dN/dS **SNFF:** “non-front-fanged” advanced snakes

doi: 10.1371/journal.pone.0081827.t001

observations lacked significant statistical support (indicated by the density and compactness of the circles). Branch-site REL identified four, three and two branches ($p < 0.05$), respectively in elapid, viperid and ‘non-front-fanged’ advanced snake NGFs as evolving under the influence of episodic diversifying selection. This highlights the episodic nature of natural selection on the evolution of advanced snake venom-specific nerve growth factors (Figure S9). This test failed to identify positive diversifying selection in other non-venomous reptilian or mammalian NGF and any BDNF and NT-3 gene lineages, including those of the advanced snakes.

Discussion

Distinct rates and trajectories of neurotrophin evolution

Evidence provided by various analyses (codeml site, branch, branch-site and clade-specific models: M8, M2a, M3, M0, two-ratio model, branch-site test A, clade model c; HyPhy: SLAC, FEL, REL, MEME, FUBAR, integrative analyses, branch-site REL; TreeSAAP: amino acid-level selection assessment and the evolutionary fingerprint analysis) not only revealed the strong influence of positive diversifying selection on Elapidae venom NGF (and to a lesser extent on the venom NGF of other advanced snake lineages, particularly the Viperidae NGF), but also highlighted the role of purifying selection pressures in shaping the molecular evolution of various non-venomous reptilian and mammalian NGF lineages (Tables 1–3; Figures 1–5; Tables S1–S4; Figures 4 and 5; Figures S6–S9). The accumulation of variation in advanced snake NGFs, in a fashion similar to other venom proteins [5,7,8] is suggestive of a role in envenomation and prey-capture. As envenomation is their primary method of prey subjugation, elapid and viperid snakes rely on the swift and potent action of venom for foraging. Hence, most elapid and viperid venom components experience extreme selection pressures, likely as a result of a co-evolutionary arms race with their prey [5,7]. Although NGF constitutes a very small proportion (1–5 mg/g of venom) of the total venom in most snakes [29,41–43], the venom of snakes of

the genus *Oxyuranus* (Taipans) has been shown to contain larger quantities of NGF [44–46]. NGF seems to make up nearly 0.5% (w/w) of the total venom injected by these snakes [44–46]. Thus, NGF could be secreted in elevated amounts by several species of caenophidian snakes, which could be enough to cause toxicity in the prey. ‘Non-front-fanged’ advanced snakes have been shown to possess several rapidly evolving venom components that participate in prey envenoming [47–52]. Selection analyses in the present study identify several residues and branches in both the ‘non-front-fanged’ and viperid advanced snake NGF lineages as evolving under positive Darwinian selection (Table 1: integrative approach; Table S4: branch-site and clade models; Figure S6, S9b and S9c). However, due to the scarcity of NGF sequences from both these lineages, the extent to which NGF participates in their envenoming remains elusive.

The aforementioned analyses also demonstrated the influence of negative selection on the evolution of all reptilian and mammalian BDNF and NT-3 gene lineages examined in this study. The lack of variation in these genes despite their origin over 300 million years ago is probably due to their significance in neuronal development, maintenance and survival. NGF, BDNF and NT-3 deficiencies have been correlated with severe peripheral neuropathy and death of the organism [53].

Focal mutagenesis of venom nerve growth factors

The synthesis and secretion of venom proteins is an energetically expensive process [54–56]. Hence, mutations that disrupt the structure/function of proteins are filtered out of the population by negative selection pressures, favouring the conservation of catalytic and structurally important core residues. As a result, most mutations in proteins can be found concentrated in structurally and/or functionally unimportant regions. Moreover, the accumulation of point mutations under the influence of positive Darwinian selection in certain regions of the toxin, such as the molecular surface (a phenomenon we refer to as focal mutagenesis), may confer adaptive

Table 2. Molecular evolution of Brain-derived Neurotrophic Factors (BDNF).

	SLAC ^a	FEL ^b	REL ^c	FUBAR ^d	Integrative	MEME ^e	M8	M2a
Elapidae								
$\omega > 1^f$	0	0	0	0	0		0	0
$\omega < 1^g$	1	3	0	3	3	0	-	-
$\omega =$	0.09	-	0.25	-	-		0.09	0.09
Viperidae								
$\omega > 1^f$	0	0	0	0	0		0	0
$\omega < 1^g$	1	3	13	4	13	0	-	-
$\omega =$	0.10	-	-	-	-		0.11	0.11
NFF								
$\omega > 1^f$	0	0	0	0	0		0	0
$\omega < 1^g$	5	19	14	22	22	0	-	-
$\omega =$	0.08	-	0.35	-	-		0.06	0.06
Typhlopoidea								
$\omega > 1^f$	0	0	0	0	0		0	0
$\omega < 1^g$	27	47	0	93	93	0	-	-
$\omega =$	0.07	-	0.11	-	-		0.07	0.07
Boidae								
$\omega > 1^f$	0	0	1	1	1		0	0
$\omega < 1^g$	3	10	3	10	10	0	-	-
$\omega =$	0.0	-	0.28	-	-		0.09	0.09
Pythonidae								
$\omega > 1^f$	0	0	0	1	1		0	0
$\omega < 1^g$	0	2	11	7	11	0	-	-
$\omega =$	0.21	-	0.23	-	-		0.15	0.15
Iguania								
$\omega > 1^f$	0	0	0	0	0		0	0
$\omega < 1^g$	24	56	111	99	24	1	-	-
$\omega =$	0.05	-	-	-	0.05		0.05	0.05
Anguimorpha								
$\omega > 1^f$	0	0	2	0	0		0	0
$\omega < 1^g$	9	22	9	23	9	0	-	-
$\omega =$	0.08	-	0.15	-	0.08		0.06	0.06
Crocodyles								
$\omega > 1^f$	0	0	0	0	0		0	0
$\omega < 1^g$	0	8	20	1	0	0	-	-
$\omega =$	0.09	-	-	-	0.09		0.12	0.12
Turtles								
$\omega > 1^f$	0	0	4	0	5		0	0
$\omega < 1^g$	6	18	2	18	18	1	-	-
$\omega =$	0.10	-	0.14	-	-		0.05	0.06
Mammals								
$\omega > 1^f$	0	0	2	0	4		0	0
$\omega < 1^g$	48	72	51	113	72	2	-	-
$\omega =$	0.05	-	-	-	-		0.05	0.06

significance. We and others have postulated that predatory venom components undergo focal mutagenesis that results in the generation of a diversity of novel amino acids (and their side-chains) on the molecular surface, which could *non-specifically* interact with prey cells and cause a myriad of pharmacological effects [11,52,57–61]. Mapping of hypermutable sites onto the three-dimensional homology model of elapid NGFs and the calculation of accessible surface

area (ASA) ratio for each residue indicated that 92% of the positively selected residues in these proteins are located on the molecular surface (Table 4; remaining sites couldn't be conclusively assigned to exposed/buried class). Moreover, a large proportion of these positively selected sites (74%) were found concentrated in the β -chain of Elapidae NGF, which is the only secreted region of the mature toxin and hence the only region that is likely to experience a coevolutionary arms race

Table 2 (continued).**a:** Single Likelihood Ancestor Counting**b:** Fixed-effects likelihood**c:** Random-effects likelihood**d:** Fast, Unconstrained Bayesian Approximation**Integrative:** Sites detected in common by SLAC, FEL, REL, FUBAR and MEME**e:** Sites detected as experiencing episodic diversifying selection (0.05 significance) by the Mixed Effects Model Evolution (MEME)**M8:** Positively-selected sites detected using the Bayes Empirical Bayes approach implemented in M8. Sites detected at 0.99 and 0.95 significance are indicated in the parenthesis**M2a:** Positively-selected sites detected using the Bayes Empirical Bayes approach implemented in M2a. Sites detected at 0.99 and 0.95 significance are indicated in the parenthesis**f:** Number of positively selected sites at 0.05 significance (for SLAC, FEL) or 50 Bayes factor (for REL) / number of sites under pervasive diversifying selection at the posterior probability ≥ 0.9 (FUBAR)**g:** Number of negatively selected sites at 0.05 significance (for SLAC, FEL) or 50 Bayes factor (for REL) / number of sites under pervasive purifying selection at the posterior probability ≥ 0.9 (FUBAR) **ω :** mean dN/dSNFF: "non-front-fanged" advanced snakes

doi: 10.1371/journal.pone.0081827.t002

(Figure S5). Thus, it is highly likely that elapid NGFs follow the regime of focal mutagenesis and favour the accumulation of hypermutable sites in specific regions of the secreted toxin that are able to influence the fitness of the animal. Rapidly evolving caenophidian NGFs with a diversity of amino acids on their molecular surface could not only induce a plethora of pharmacological effects by non-specifically interacting with novel receptors of the prey, but could further aid in evading the prey immune response [62].

Putative role of nerve growth factors in venom

Nerve growth factors are known to inhibit venom metalloproteinase-dependent proteolysis [63]. Hence, the presence of large amounts of NGF, relative to metalloproteinases in viper venom [64,65] has been viewed as an indication of their role in prevention of venom autolysis [63]. However, additional plausible mechanisms of action exist for both the direct and indirect participation of rapidly evolving caenophidian NGF in prey envenoming, potentially resulting in an increase in the potency of the toxicoferan venom arsenal. The injection of large amounts of NGF into the lymphatic system of prey would result in NGF overdose, triggering a variety of reactions. Nerve growth factors are also known to cause apoptosis of cells lacking the expression of TrkA receptors [66–68] and thus the caenophidian venom NGFs could act as cytotoxic proapoptotic factors, after being introduced into tissues where they are not typically found. Venom components like hyaluronidase have been shown to degrade hyaluronic acid in the extra-cellular matrix and function as spreading factors, facilitating the easier diffusion of other venom components [69]. NGF, which is known to trigger the release of granules containing histamines, serotonin and other chemical mediators from mastocytes [70,71], could perform a similar ancillary venom function by facilitating the efficient absorption of other venom components. The degranulation of mast cells that are located primarily in perivascular spaces, often close to neurons and blood vessels, has been correlated

with increased vascular permeability and even neurogenic inflammations [72]. While controlled degranulation aids in orchestrating and mounting acute inflammatory reactions, massive releases can be associated with anaphylaxis, bronchoconstriction (suffocation) and vasodilation [73]. Over the years, both glycosylated and non-glycosylated forms of NGFs have been reported from snake venoms [44,74–76]. Glycosylation has been proposed to prolong the circulation time of serum proteins by increasing their stability [77]. Interestingly, Asn 23, which has been postulated as a putative glycosylation site in mature venom NGF [78,79], is highly conserved in the entire squamate lineage (except in a few elapid snakes: genus *Naja*, *Notechis scutatus*, *Pseudechis australis*, *Pseudechis porphyriacus*, etc.), while being completely absent from mammalian and turtle NGFs (Figure S5). These observations are not surprising since mammalian NGFs are not known to undergo glycosylation. Snake venom NGFs could undergo posttranslational glycosylation in order to circulate longer in the bloodstream of the prey and spread further throughout the prey animal's system. Thus, caenophidian NGFs could directly and/or indirectly participate in prey-envenoming through a number of plausible mechanisms.

Putative duplication of nerve growth factors in *Elapidae*

Genes encoding venom proteins are known to evolve via the birth-and-death model of evolution where new genes are created by repeated duplication events, and are subsequently either maintained in the genome, deleted or become non-functional pseudogenes [7,8]. The molecular phylogeny of NGFs clearly indicates the presence of two independent NGF genes in *Naja sputatrix*, with very strong support for the node separating them (915 bootstrap and 1.0 Bayesian posterior probability: Figure 1). The phylogenetic placement of these two gene homologues suggests that the gene duplication event responsible for their origin occurred in an early elapid ancestor (Figure 1) – we might therefore expect to find additional copies

Table 3. Molecular evolution of of Neurotrophin 3 (NT-3).

	SLAC ^a	FEL ^b	REL ^c	FUBAR ^d	Integrative	MEME ^e	M8	M2a
Elapidae								
$\omega > 1^f$	0	0	0	0	0		0	0
$\omega < 1^g$	1	4	0	4	4	0	-	-
$\omega =$	0.19	-	0.21	-	-		0.18	0.19
Viperidae								
$\omega > 1^f$	0	0	0	0	0		0	0
$\omega < 1^g$	0	5	0	3	5	0	-	-
$\omega =$	0.30	-	0.49	-	-		0.28	0.30
NFF								
$\omega > 1^f$	0	2	5	1	8		0	0
$\omega < 1^g$	43	61	79	82	95	2	-	-
$\omega =$	0.22	-	0.35	-	-		0.20	0.27
Typhlopoidea								
$\omega > 1^f$	1	2	8	2	10		0	0
$\omega < 1^g$	48	69	49	106	106	2	-	-
$\omega =$	0.24	-	0.49	-	-		0.23	0.26
Boidae								
$\omega > 1^f$	0	0	0	1	3		0	0
$\omega < 1^g$	7	23	10	26	26	2	-	-
$\omega =$	0.24	-	0.51	-	-		0.24	0.28
Scinciformata								
$\omega > 1^f$	0	0	0	0	0		0	0
$\omega < 1^g$	7	27	75	30	75	0	-	-
$\omega =$	0.13	-	-	-	-		0.12	0.13
Iguania								
$\omega > 1^f$	0	2	4	3	7		0	0
$\omega < 1^g$	27	41	23	58	58	3	-	-
$\omega =$	0.25	-	0.43	-	-		0.23	0.26
Anguimorpha								
$\omega > 1^f$	0	0	0	0	1		0	0
$\omega < 1^g$	7	20	All	23	23	1	-	-
$\omega =$	0.24	-	0.32	-	-		0.21	0.24
Crocodyles								
$\omega > 1^f$	0	0	0	0	0		0	0
$\omega < 1^g$	0	1	all	1	1	0	-	-
$\omega =$	0.08	-	0.08	-	-		0.07	0.07
Turtles								
$\omega > 1^f$	0	0	0	0	0		0	0
$\omega < 1^g$	3	4	all	5	5	0	-	-
$\omega =$	0.19	-	0.21	-	-		0.17	0.17
Mammals								
$\omega > 1^f$	0	0	1	0	2		0	0
$\omega < 1^g$	77	124	134	167	172	1	-	-
$\omega =$	0.06	-	-	-	-		0.05	0.07

of NGF in other elapid snakes in the future. Although one might initially assume that the original copy of NGF would continue the ancestral role in homeostasis, thereby releasing the novel duplicate gene to evolve a function specific to venom, in *N. sputatrix* both NGF genes were found to be expressed in the venom gland [80]. However, considering how important NGF and other neurotrophins appear to be for the survival of an organism [53], it is highly unlikely that both gene copies would

be exclusively associated with a role in prey envenoming. We therefore propose that co-expression of an NGF gene occurs in different tissue types, similar to that recently proposed for other toxin families [81]. Specifically, we postulate that despite the elevated mutation rate in the Elapidae NGF, the ancestral copy of the gene has retained its biochemical functions by undergoing focal mutagenesis and continues to ensure homeostasis. A greater accumulation of mutations on the

Table 3 (continued).

a: Single Likelihood Ancestor Counting
b: Fixed-effects likelihood
c: Random-effects likelihood
d: Fast, Unconstrained Bayesian AppRoximation
Integrative: Sites detected in common by SLAC, FEL, REL, FUBAR and MEME
e: Sites detected as experiencing episodic diversifying selection (0.05 significance) by the Mixed Effects Model Evolution (MEME)
M8: Positively-selected sites detected using the Bayes Empirical Bayes approach implemented in M8. Sites detected at 0.99 and 0.95 significance are indicated in the parenthesis
M2a: Positively-selected sites detected using the Bayes Empirical Bayes approach implemented in M2a. Sites detected at 0.99 and 0.95 significance are indicated in the parenthesis
f: Number of positively selected sites at 0.05 significance (for SLAC, FEL) or 50 Bayes factor (for REL) / number of sites under pervasive diversifying selection at the posterior probability ≥ 0.9 (FUBAR)
g: Number of negatively selected sites at 0.05 significance (for SLAC, FEL) or 50 Bayes factor (for REL) / number of sites under pervasive purifying selection at the posterior probability ≥ 0.9 (FUBAR)
 ω : mean dN/dSNFF: "non-front-fanged" advanced snakes
doi: 10.1371/journal.pone.0081827.t003

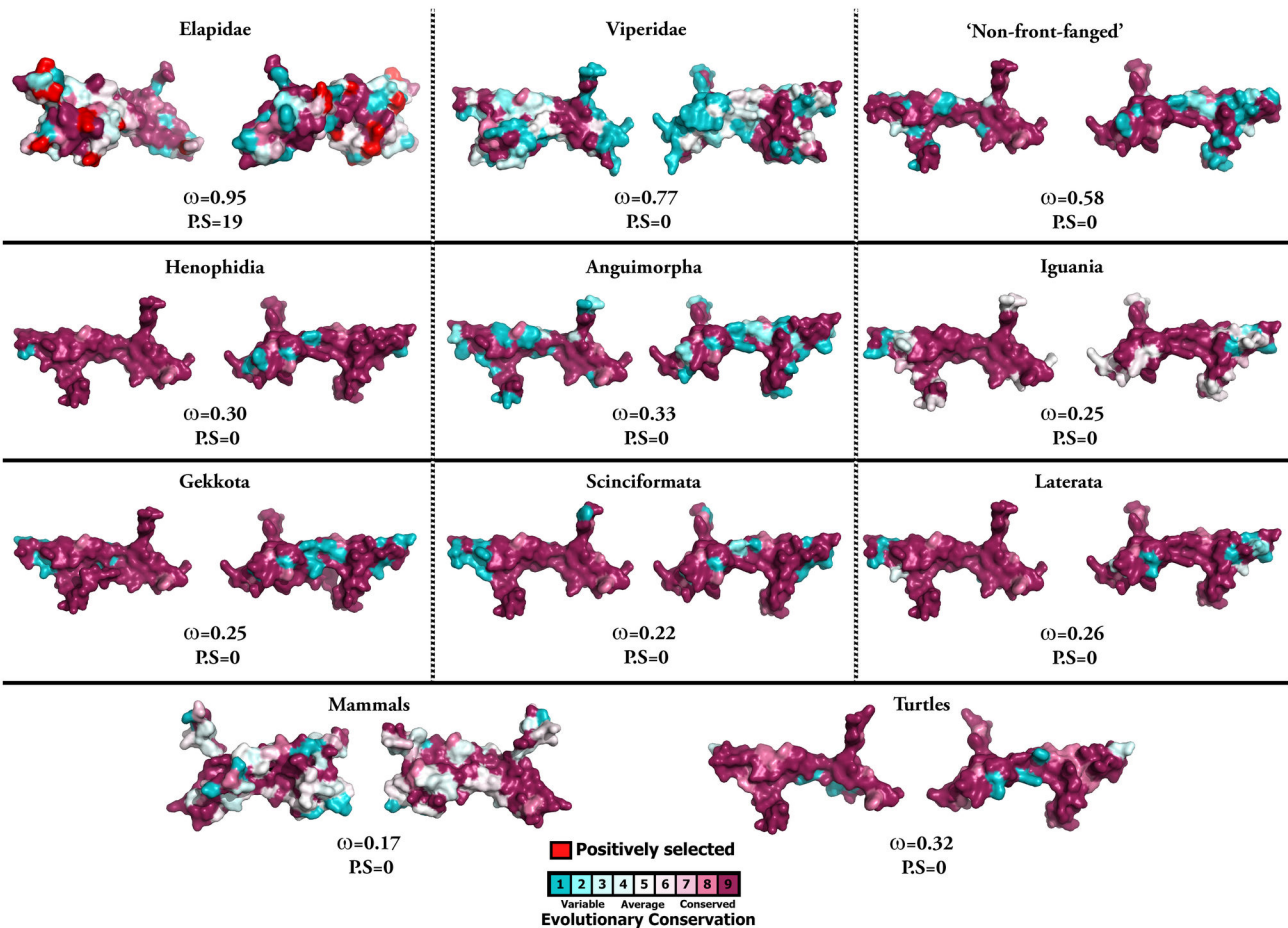


Figure 4. Molecular evolution of nerve growth factors. Three-dimensional homology models of nerve growth factors, depicting the locations of positively selected sites (in red). The ω values (Model 8) for the respective taxa along with the number of positively selected sites (PP ≥ 0.95 , Bayes-Empirical Bayes approach) are indicated.

doi: 10.1371/journal.pone.0081827.g004

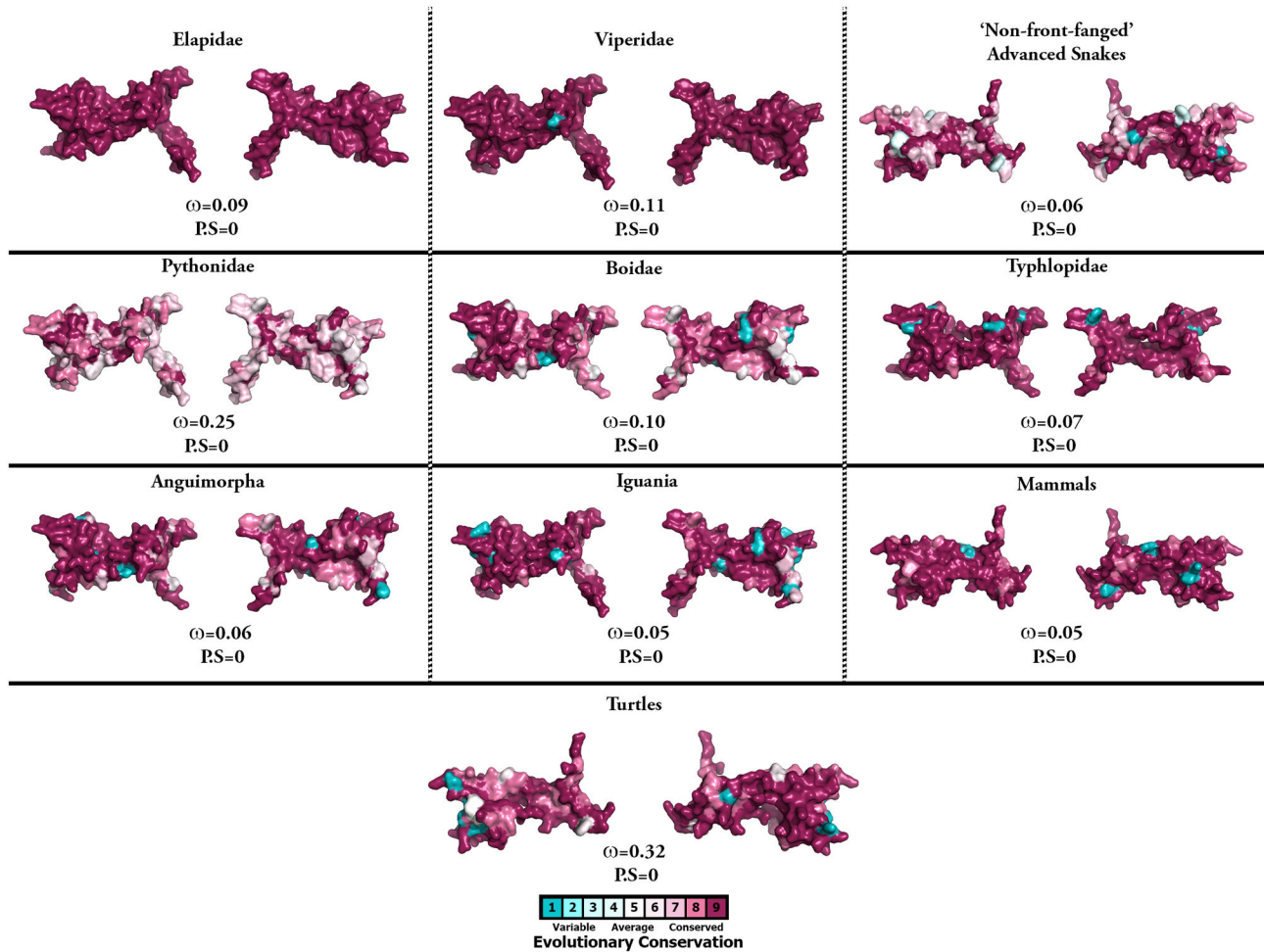


Figure 5. Molecular evolution of brain-derived neurotrophic factors. Three-dimensional homology models of brain-derived neurotrophic factors (BDNF), depicting the molecular evolution of these non-venomous homologues. The ω values (Model 8) for the respective taxa (PP ≥ 0.95 , Bayes-Empirical Bayes approach) are indicated.

doi: 10.1371/journal.pone.0081827.g005

molecular surface and the loops of the protein ensures the preservation of structurally and functionally important core residues. Clearly, a large number of residues across the entire vertebrate NGF lineage appear to be highly conserved (Figure S5). Moreover, it has been demonstrated on numerous occasions that snake venom NGFs exhibit biological activities similar to their mammalian counterparts [44,74]. This would seemingly free the duplicate elapid NGF gene to participate exclusively in prey envenoming. In contrast to the Elapidae NGF, only one copy was found in the venom glands of the Anguimorpha lizard *Abronia graminea*, which was identical to the one recovered from genomic sequencing for use in taxonomy studies [27]. Moreover, despite extensive BLAST searches using sequence templates from a wide array of mammalian and reptilian NGFs against the genomes of *Anolis carolinensis* (Iguania) and *Python molurus* (Henophidia), we were only able to retrieve a single copy of the NGF gene. Hence, we further speculate that the ancestral NGF gene is co-

expressed in physiological tissues and the venom-gland, while the new gene duplicates are more likely to be tissue-specifically expressed in the venom gland alone. Experimental investigation in the future regarding the copy number of NGF should reveal if the NGF duplication occurred in an early elapid ancestral lineage or at the base of the Toxicofera phylogenetic tree.

Note on the usage of NGF in the phylogenetic reconstructions of Toxicofera reptiles

NGF is a commonly used gene in reptilian phylogenetic studies [82–85]. However, the detection of gene duplication events and the accelerated evolution of the advanced snake NGFs indicate that these genes should not be employed in phylogenetic inferences of Toxicofera reptiles, as these factors can undermine such analyses [86].

Table 4. Nucleotide and complementary protein-level selection assessment of Elapidae Nerve Growth Factors (NGF).

Codon	Amino Acid	M2a ^a	M8 ^b	Property ^c	Magnitude ^d	ASA
57	R	4.241±0.515 (0.994)	3.894±0.510 (0.998)*	-	-	-
72	A	4.260±0.453 (0.999)**	3.899±0.493 (1.0)**	-	-	-
116	T	4.227±0.553 (0.990)*	3.892±0.513 (0.997)**	R_{α}, α_m	6, 7	-
123	A	4.243±0.509 (0.994)	3.895±0.506 (0.998)*	R_{α}, α_m	6, 7	-
125	R	4.132±0.769 (0.961)*	3.857±0.601 (0.986)*	-	-	-
127	N	4.262±0.448 (1.0)**	3.900±0.492 (1.0)**	R_{α}, α_m	6, 7	-
133	Q	4.262±0.446 (1.0)**	3.900±0.491 (1.0)**	R_{α}, α_m	6, 7	-
144	D	4.258±0.459 (0.999)**	3.899±0.495 (1.0)**	R_{α}, α_m	6, 7	69.5 Exposed
150	T	4.252±0.481 (0.997)**	3.897±0.500 (0.999)**	α_m	6	77.5 Exposed
155	R	4.146±0.740 (0.965)*	3.867±0.576 (0.990)*	α_m	6	87.0 Exposed
158	V	4.247±0.498 (0.995)**	3.896±0.504 (0.999)**	-	-	76.9 Exposed
163	E	4.257±0.464 (0.998)*	3.898±0.496 (0.999)**	α_m	6	65.9 Exposed
167	L	4.024±0.940 (0.929)	3.824±0.670 (0.976)*	α_m	6	65.0 Exposed
170	E	4.260±0.453 (0.999)**	3.900±0.493 (1.0)**	-	-	73.0 Exposed
182	R	3.951±1.036 (0.907)	3.789±0.736 (0.965)*	-	-	57.3 Exposed
207	Q	4.255±0.471 (0.998)**	3.898±0.499 0.999**	R_{α}	6	32.2 NA
211	R	4.260±0.455 (0.999)*	3.899±0.493 (1.0)**	R_{α}	6	53.1 Exposed
219	Q	4.262±0.446 (1.0)**	3.900±0.491 (1.0)**	R_{α}	6	73.8 Exposed
240	D	4.004±0.970 (0.923)	3.804±0.712 (0.969)*	R_{α}	6	66.9 Exposed
167	L	4.024±0.940 (0.929)	3.824±0.670 (0.976)*	α_m	6	65.0 Exposed
170	E	4.260±0.453 (0.999)**	3.900±0.493 (1.0)**	-	-	73.0 Exposed
182	R	3.951±1.036 (0.907)	3.789±0.736 (0.965)*	-	-	57.3 Exposed
207	Q	4.255±0.471 (0.998)**	3.898±0.499 0.999**	R_{α}	6	32.2 NA
211	R	4.260±0.455 (0.999)*	3.899±0.493 (1.0)**	R_{α}	6	53.1 Exposed
219	Q	4.262±0.446 (1.0)**	3.900±0.491 (1.0)**	R_{α}	6	73.8 Exposed
240	D	4.004±0.970 (0.923)	3.804±0.712 (0.969)*	R_{α}	6	66.9 Exposed

Table 4 (continued).

Amino-acid property symbols used: Power to be in the middle of α -helix (α_m), Solvent accessible reduction ratio (R_α)

PAML

a, b: Bayes Empirical Bayes (BEB) posterior probability and post-mean omega (indicated in brackets) for the sites detected as positively selected by the site models M2a and M8, respectively. Sites detected as positively selected at 0.95 and 0.99 posterior probability by the Bayes Empirical Bayes approach of M8 are represented by * and **, respectively.

TreeSAAP

c: amino acid property experiencing positive diversifying selection

d: magnitude of selection on the amino acid property

ASA: Accessible surface area.

Note: Codon sites with significant support from both nucleotide and protein-level selection analyses are highlighted in bold.

doi: 10.1371/journal.pone.0081827.t004

Conclusions

In this study, we show that BDNF and NT-3, two major members of the neurotrophin family, lack variations completely and evolve under the regime of negative selection, indicating their extreme importance in the maintenance of homeostasis. In complete contrast, this study points for the first time towards venom-specific NGF evolving in a manner typical of toxic proteins and consequently suggests a hitherto unappreciated role in envenomation. We highlight an exquisite mechanism of venom evolution where focal mutagenesis has facilitated the transformation of a key regulatory protein into a toxin, without hampering its participation in homeostasis. Structure-function studies of mutant NGF may reveal a suite of novel activities, some of which may be of use in drug design and development.

Materials and Methods

Phylogenetic Analyses

Phylogenetic analyses were performed to allow the reconstruction of the molecular evolutionary history of vertebrate NGF, BDNF and NT-3 genes. Nucleotide sequences were downloaded from the National Center for Biotechnology Information database (NCBI: <http://www.ncbi.nlm.nih.gov/>) and a dataset of 1183 sequences (NGF: 308; BDNF: 434; NT-3: 441) was compiled. Accession numbers of all sequences analysed in this study are available in Tables S5.1-S5.3. Unpublished sequences are available in the File S1. Resultant sequence sets were aligned using MUSCLE [87]. The phylogenetic relationships were determined using Bayesian and maximum-likelihood approaches. MrBayes version 3.2 [88] was used for Bayesian inference. Tree searches were run using four Markov chains for a minimum of 10 million generations, sampling every 100th tree. The log likelihood score of each saved tree was plotted against the number of generations to establish the point at which the log-likelihood scores of the analyses reached their asymptote. 25% of the total trees sampled were conservatively discarded as burnin. The posterior probabilities for clades were established by constructing a majority rule consensus tree for all trees generated after the completion of the burnin. The analyses were repeated three times to make sure that the trees

generated were not clustered around local optima. An optimal maximum likelihood phylogenetic tree was obtained using PhyML 3.0 [89] and node support was evaluated with 1,000 bootstrapping replicates.

Test for Recombination

To overcome the effects of recombination on the phylogenetic and evolutionary interpretations [90], we employed Single Breakpoint algorithm implemented in the HyPhy package and assessed recombination on all the toxin forms examined in this study [91,92]. When potential breakpoints were detected using the small sample Akaike information Criterion (AICc), the sequences were compartmentalized before conducting selection analyses, so as to allow recombining units to have distinct phylogeny.

Selection Analyses

We evaluated the influence of natural selection on various members of vertebrate neurotrophins using maximum-likelihood models [93,94] implemented in CODEML of the PAML [95]. In order to detect evolutionary selection pressures acting upon individual lineages, we employed the two-ratio model as well as the optimized branch-site test [96,97]. The two-ratio model evaluates selection across the lineages alone, while the branch-site model allows ω to vary across the sites of the protein and along the branches in the tree. The latter is known for its reasonable power and accuracy to detect short bursts of episodic adaptations [97]. However, both the two-ratio and branch-site models require the foreground (lineages under positive selection) and background lineages (lineages lacking positive selection) to be defined *a priori*. Such predefined biological hypotheses are often unavailable and it becomes difficult to define the foreground lineages. Therefore, we treated each clade being compared as foreground branches alternatively and tested multiple hypotheses. A likelihood-ratio test was then conducted by comparing the model that allows ω to be greater than 1 in the foreground branch, with a null model that does not (ω constrained 1). It has been suggested that while implementing multiple hypotheses using branch and branch-site models, it is necessary to control the family-wise error rate (FWER or Type I error) [97]. We used Bonferroni correction to account for such errors. It uses α/n as the

significance level to test each hypothesis; where ' α ' is the significance level and ' n ' is the number of independent true null hypotheses. We further utilized branch-site REL implemented in HyPhy to identify lineages affected by positive selection pressures. Unlike the aforementioned lineage-specific models, branch-site REL does not require the identification of foreground and background branches *a priori*.

Because lineage-specific models either assume a single ω for the entire tree or assess the effects of selection only along the branches, they often fail to identify regions in proteins that might be affected by episodic selection pressures and ultimately underestimate the strength of selection. Hence, we employed site-specific models which estimate positive selection statistically as a non-synonymous-to-synonymous nucleotide-substitution rate ratio (ω) significantly greater than 1. We compared likelihood values for three pairs of models with different assumed ω distributions as no *a priori* expectation exists for the same: M0 (constant ω rates across all sites) versus M3 (allows the ω to vary across sites within ' n ' discrete categories, $n \geq 3$); M1a (a model of neutral evolution) where all sites are assumed to be either under negative ($\omega < 1$) or neutral selection ($\omega = 1$) versus M2a (a model of positive selection) which in addition to the site classes mentioned for M1a, assumes a third category of sites; sites with $\omega > 1$ (positive selection) and M7 (β) versus M8 (β and ω), and models that mirror the evolutionary constraints of M1 and M2 but assume that ω values are drawn from a β distribution [98]. Only if the alternative models (M3, M2a and M8: allow sites with $\omega > 1$) show a better fit in Likelihood Ratio Test (LRT) relative to their null models (M0, M1a and M7: do not allow sites $\omega > 1$), are their results considered significant. LRT is estimated as twice the difference in maximum likelihood values between nested models and compared with the χ^2 distribution with the appropriate degree of freedom (i.e., the difference in the number of parameters between the two models). The Bayes empirical Bayes (BEB) approach [97] was used to identify amino acids under positive selection by calculating the posterior probabilities that a particular amino acid belongs to a given selection class (neutral, conserved or highly variable). Sites with greater posterior probability (PP $\geq 95\%$) of belonging to the ' $\omega > 1$ class' were inferred to be positively selected.

SLAC, FEL, REL and FUBAR [99,100] implemented in HyPhy [101] were employed to provide additional support to the aforementioned analyses and to detect sites evolving under the influence of positive and negative selection. MEME [102] was also used to detect episodic diversifying selection. Further support for the results of the nucleotide-level selection analyses was obtained and the radicalness of mutations were assessed using a complementary protein-level approach implemented in TreeSAAP [103].

Direct comparison of ω values computed from different datasets can be misleading, as they can have different proportion of sites under selection. Hence, we assessed the selection pressures shaping the evolution of NGF along various toxicoforan reptilian lineages by employing clade model analyses implemented in codeml and simultaneously estimated ω values [104]. The significance of the analysis was tested by comparing the likelihood of this model with that of model M1a.

To clearly depict the proportion of sites under different regimes of selection, an evolutionary fingerprint analysis was carried out using the evolutionary selection distance (ESD) algorithm implemented in datamonkey [91,105,106]. Evolutionary fingerprint analysis fits a versatile general discrete bivariate model of site-to-site variation in selection pressures and comprises a description of the number of selective classes, the dN/dS (ω) ratio for each class and the assignment of sites to classes [91,106].

Structural Analyses

To depict the natural selection pressures influencing the evolution of various neurotrophins, we mapped the sites under positive selection on the homology models created using Phyre 2 webserver [107]. Pymol 1.3 [108] was used to visualize and generate the images of homology models. The consurf webserver [109] was used for mapping the evolutionary selection pressures on the three-dimensional homology models. GETAREA [110] was used to calculate the ASA ratio or the solvent exposure of amino-acid side chains. It uses the atom co-ordinates of the PDB file and indicates if a residue is buried or exposed to the surrounding medium by comparing the ratio between side-chain ASA and the "random coil" values per residue. An amino-acid is considered to be buried if ASA is less than 20% and exposed if ASA is $\geq 50\%$.

Supporting Information

Table S1. (S1.1 – S1.11) Details of selection analyses of nerve growth factors (NGF). a: dn/ds (weighted average). b: Significance of the model in comparison with the null model. c: Number of sites with $\omega > 1$ under the Bayes empirical Bayes approach with a posterior probability (PP) more than or equal to 0.99 and 0.95. * Models which allow $\omega > 1$. (PDF)

Table S2. (S2.1 – S2.11) Details of selection analyses of brain-derived neurotrophic factors (BDNF). a: dn/ds (weighted average). b: Significance of the model in comparison with the null model. c: Number of sites with $\omega > 1$ under the Bayes empirical Bayes approach with a posterior probability (PP) more than or equal to 0.99 and 0.95. * Models which allow $\omega > 1$. (PDF)

Table S3. (S3.1-S3.11) Details of selection analyses of neurotrophin-3 (NT3). a: dn/ds (weighted average). b: Significance of the model in comparison with the null model. c: Number of sites with $\omega > 1$ under the Bayes empirical Bayes approach with a posterior probability (PP) more than or equal to 0.99 and 0.95. * Models which allow $\omega > 1$. (PDF)

Table S4. Lineage-specific selection analyses of nerve growth factors (NGF). a: dn/ds (weighted average). b: Significance of the model in comparison with the null model. * Significant after Bonferroni correction. NS: Not significant.

Significantly detected positively selected lineages are highlighted in bold.
(PDF)

Table S5. (S5.1-S5.3) Sequences analysed.
(PDF)

Figure S1. Maximum-likelihood molecular phylogeny of nerve growth factors (NGF). Branches with bootstrap support of less than 850 (out of 1000 bootstrap replicates) are highlighted in grey. [NFF: “non-front-fanged” advanced snakes; Atr: Atractaspidae; Sci: Scinciformata].
(PDF)

Figure S2. Maximum-likelihood molecular phylogeny of brain-derived neurotrophic factors (BDNF). Branches with bootstrap support of less than 850 (out of 1000 bootstrap replicates) are highlighted in grey.
(PDF)

Figure S3. Maximum-likelihood molecular phylogeny of neurotrophin-3 (NT3). Branches with bootstrap support of less than 850 (out of 1000 bootstrap replicates) are highlighted in grey.
(PDF)

Figure S4. Maximum-likelihood phylogeny of neurotrophins. Branches with bootstrap support of less than 850 (out of 1000 bootstrap replicates) are highlighted in grey. [NFF: ‘non-front-fanged’ advanced snakes; Pyt: Pythonidae; Mam: Mammals].
(PDF)

Figure S5. Alignment of vertebrate nerve growth factors.
(PDF)

Figure S6. Evolutionary fingerprint of nerve growth factors (NGF). Estimates of the distribution of synonymous (a) and non-synonymous (b) substitution rates inferred for various reptilian and mammalian nerve growth factor (NGF) lineages are shown here. The ellipses reflect a Gaussian-approximated variance in each individual rate estimate, and coloured pixels show the density of the posterior sample of the distribution for a given rate. The diagonal line represents the idealized neutral evolution regime ($\omega = 1$), points above and below the line correspond to positive selection ($\omega > 1$) and negative selection ($\omega < 1$), respectively. Site model 8 omega (ω) along with the total number of positively selected sites detected by its Bayes Empirical Bayes (BEB) approach are also indicated below.
(PDF)

Figure S7. Evolutionary fingerprint of brain-derived neurotrophic factors (BDNF). Estimates of the distribution of

synonymous (a) and non-synonymous (b) substitution rates inferred for various reptilian and mammalian brain-derived neurotrophic factor (BDNF) lineages are shown here. The ellipses reflect a Gaussian-approximated variance in each individual rate estimate, and coloured pixels show the density of the posterior sample of the distribution for a given rate. The diagonal line represents the idealized neutral evolution regime ($\omega = 1$), points above and below the line correspond to positive selection ($\omega > 1$) and negative selection ($\omega < 1$), respectively. Site model 8 omega (ω) along with the total number of positively selected sites detected by its Bayes Empirical Bayes (BEB) approach are also indicated below.
(PDF)

Figure S8. Evolutionary fingerprint of neurotrophin-3 (NT3). Estimates of the distribution of synonymous (a) and non-synonymous (b) substitution rates inferred for various reptilian and mammalian neurotrophin-3 (NT-3) lineages are shown here. The ellipses reflect a Gaussian-approximated variance in each individual rate estimate, and coloured pixels show the density of the posterior sample of the distribution for a given rate. The diagonal line represents the idealized neutral evolution regime ($\omega = 1$), points above and below the line correspond to positive selection ($\omega > 1$) and negative selection ($\omega < 1$), respectively. Site model 8 omega (ω) along with the total number of positively selected sites detected by its Bayes Empirical Bayes (BEB) approach is also indicated below.
(PDF)

Figure S9. Branch-site REL: Caenophidian nerve growth factors (NGF). The hue of each colour indicates strength of selection, with primary red corresponding to $\omega > 5$, primary blue to $\omega = 0$ and grey to $\omega = 1$. The width of each colour component represents the proportion of sites in the corresponding class. Thicker branches have been classified as undergoing episodic diversifying selection (indicated by arrows) by the sequential likelihood ratio test at corrected $p \leq 0.05$.
(PDF)

File S1. Unpublished sequences.
(DOCX)

Acknowledgements

KS is thankful to Dr. Rohan Kolla and Dr. Krishna Mohan for the useful discussions related to the pathology of envenoming.

Author Contributions

Conceived and designed the experiments: BGF NRC AA KS. Performed the experiments: TJ EU SA KS. Analyzed the data: NV BGF NRC KS. Contributed reagents/materials/analysis tools: AA BGF VV GFK NV KV. Wrote the manuscript: KS BGF TJ NRC EU NV SA GFK KV VV AA.

References

- Fry BG, Roelants K, Champagne DE, Scheib H, Tyndall JDA et al. (2009) The toxicogenomic multiverse: Convergent recruitment of proteins into animal venoms. *Annu Rev Genomics Hum Genet* 10: 483–511. doi:10.1146/annurev.genom.9.081307.164356. PubMed: 19640225.
- Bush SP, Green SM, Moynihan JA, Hayes WK, Cardwell MD (2002) Crotalidae polyvalent immune fab (ovine) antivenom is efficacious for envenomations by southern pacific rattlesnakes (*Crotalus helleri*). *Ann Emerg Med* 40: 619–624. doi:10.1067/mem.2002.129939. PubMed: 12447339.
- Heatwole H, Powell J (1998) Resistance of eels (*Gymnothorax*) to the venom of sea kraits (*Laticauda Colubrina*): a test of coevolution. *Toxicon* 36: 619–625. doi:10.1016/S0041-0101(97)00081-0. PubMed: 9643474.
- Massey DJ, Calvete JJ, Sánchez EE, Sanz L, Richards K et al. (2012) Venom variability and envenoming severity outcomes of the *Crotalus scutulatus* scutulatus (Mojave rattlesnake) from Southern Arizona. *J Proteomics* 75: 2576–2587. doi:10.1016/j.jprot.2012.02.035. PubMed: 22446891.
- Casewell NR, Wüster W, Vonk FJ, Harrison RA, Fry BG (2013) Complex cocktails: the evolutionary novelty of venoms. *Trends Ecol Evol* 28: 219–229. doi:10.1016/j.tree.2012.10.020. PubMed: 23219381.
- Olivera BM (2002) CONUS VENOM PEPTIDES: Reflections from the Biology of Clades and Species. *Annu Rev Ecol Syst* 33: 25–47. doi:10.1146/annurev.ecolsys.33.010802.150424.
- Fry BG, Wüster W, Kini RM, Brusich V, Khan A et al. (2003) Molecular evolution and phylogeny of elapid snake venom three-finger toxins. *J Mol Evol* 57: 110–129. doi:10.1007/s00239-003-2461-2. PubMed: 12962311.
- Fry BG (2005) From genome to “venome”: Molecular origin and evolution of the snake venom proteome inferred from phylogenetic analysis of toxin sequences and related body proteins. *Genome Res* 15: 403–420. doi:10.1101/gr.3228405. PubMed: 15741511.
- Escoubas P (2006) Molecular diversification in spider venoms: a web of combinatorial peptide libraries. *Mol Divers* 10: 545–554. doi:10.1007/s11030-006-9050-4. PubMed: 17096075.
- Weinberger H, Moran Y, Gordon D, Turkov M, Kahn R et al. (2010) Positions under positive selection—key for selectivity and potency of scorpion alpha-toxins. *Mol Biol Evol* 27: 1025–1034. doi:10.1093/molbev/msp310. PubMed: 20018978.
- Casewell NR, Wagstaff SC, Harrison RA, Renjifo C, Wüster W (2011) Domain loss facilitates accelerated evolution and neofunctionalization of duplicate snake venom metalloproteinase toxin genes. *Mol Biol Evol* 28: 2637–2649. doi:10.1093/molbev/msr091. PubMed: 21478373.
- Chang D, Duda TF (2012) Extensive and continuous duplication facilitates rapid evolution and diversification of gene families. *Mol Biol Evol* 29: 2019–2029. doi:10.1093/molbev/mss068. PubMed: 22337864.
- Kozminsky-Atias A, Zilberberg N (2012) Molding the business end of neurotoxins by diversifying evolution. *FASEB J* 26: 576–586. doi:10.1096/fj.11-187179. PubMed: 22009937.
- Zhu S, Peigneur S, Gao B, Lu X, Cao C et al. (2012) Evolutionary diversification of Mesobuthus α -scorpion toxins affecting sodium channels. *Mol Cell Proteomics* 11: M111: 012054. doi:10.1074/mcp.M111.012054. PubMed: 21969612.
- Fry BG, Undheim EAB, Ali SA, Jackson TNW, Debono J et al. (2013) Squeezers and Leaf-cutters: Differential Diversification and Degeneration of the Venom System in Toxicofera Reptiles. *Mol Cell Proteomics* 12: 1881–1899. doi:10.1074/mcp.M112.023143. PubMed: 23547263.
- Boldrini-França J, Corrêa-Netto C, Silva MMS, Rodrigues RS, De La Torre P et al. (2010) Snake venomomics and antivenomics of *Crotalus durissus* subspecies from Brazil: assessment of geographic variation and its implication on snakebite management. *J Proteomics* 73: 1758–1776. doi:10.1016/j.jprot.2010.06.001. PubMed: 20542151.
- Calvete JJ, Escolano J, Sanz L (2007) Snake venomomics of *Bitis* species reveals large intragenous venom toxin composition variation: application to taxonomy of congeneric taxa. *J Proteome Res* 6: 2732–2745. doi:10.1021/pr0701714. PubMed: 17559253.
- Casewell NR, Harrison RA, Wüster W, Wagstaff SC (2009) Comparative venom gland transcriptome surveys of the saw-scaled vipers (*Viperidae: Echis*) reveal substantial intra-family gene diversity and novel venom transcripts. *BMC Genomics* 10: 564. doi:10.1186/1471-2164-10-564. PubMed: 19948012.
- Fry BG, Wickramaratna JC, Hodgson WC, Alewood PF, Kini RM et al. (2002) Electrospray liquid chromatography/mass spectrometry fingerprinting of *Acanthophis* (death adder) venoms: taxonomic and toxicological implications. *Rapid Commun Mass Spectrom* 16: 600–608. doi:10.1002/rcm.613. PubMed: 11870898.
- Fry BG, Wüster W, Ryan Ramjan SF, Jackson T, Martelli P et al. (2003) Analysis of Colubroidea snake venoms by liquid chromatography with mass spectrometry: evolutionary and toxicological implications. *Rapid Commun Mass Spectrom* 17: 2047–2062. doi:10.1002/rcm.1148. PubMed: 12955733.
- Gutiérrez JM, Lomonte B, León G, Alape-Girón A, Flores-Díaz M et al. (2009) Snake venomomics and antivenomics: Proteomic tools in the design and control of antivenoms for the treatment of snakebite envenoming. *J Proteomics* 72: 165–182. doi:10.1016/j.jprot.2009.01.008. PubMed: 19344652.
- Terrat Y, Sunagar K, Fry BG, Jackson TN, Scheib H et al. (2013) *Atractaspis aterrima* Toxins: The First Insight into the Molecular Evolution of Venom in Side-Stabbers. *Toxins (Basel)* 5: 1948–1964. doi:10.3390/toxins5111948. PubMed: 24169588.
- Fry BG, Scheib H, van der Weerd L, Young B, McNaughtan J et al. (2009) Evolution of an arsenal: structural and functional diversification of the venom system in the advanced snakes (Caenophidia). *Mol Cell Proteomics* 7: 215–246. doi:10.1074/mcp.M700094-MCP200. PubMed: 17855442.
- Fry BG, Vidal N, Norman JA, Vonk FJ, Scheib H et al. (2006) Early evolution of the venom system in lizards and snakes. *Nature* 439: 584–588. doi:10.1038/nature04328. PubMed: 16292255.
- Fry BG, Winter K, Norman JA, Roelants K, Nabuurs RJA et al. (2010) Functional and structural diversification of the Anguimorpha lizard venom system. *Mol Cell Proteomics* 9: 2369–2390. doi:10.1074/mcp.M110.001370. PubMed: 20631207.
- Hogue-Angeletti RA, Frazier WA, Jacobs JW, Niall HD, Bradshaw RA (1976) Purification, characterization, and partial amino acid sequence of nerve growth factor from cobra venom. *Biochemistry* 15: 26–34. doi:10.1021/bi00646a005. PubMed: 1247508.
- Koludarov I, Sunagar K, Undheim EA, Jackson TN, Ruder T et al. (2012) Structural and molecular diversification of the Anguimorpha lizard mandibular venom gland system in the arboreal species *Abronia graminea*. *J Mol Evol* 75: 168–183. doi:10.1007/s00239-012-9529-9.
- Oda T, Ohta M, Inoue S, Ikeda K, Furukawa S et al. (1989) Amino acid sequence of nerve growth factor purified from the venom of the Formosan cobra *Naja naja atra*. *Biochem Int* 19: 909–917. PubMed: 2619756.
- Siigur E, Neuman T, Järve V, Tara A, Siigur J (1985) Isolation and characterization of nerve growth factor from *Vipera lebetina* (snake) venom. *Comp Biochem Physiol B* 81: 211–215. PubMed: 4017541.
- Hamburger V, Levi-Montalcini R (1949) Proliferation, differentiation and degeneration in the spinal ganglia of the chick embryo under normal and experimental conditions. *J Exp Zool* 111: 457–501. doi:10.1002/jez.1401110308. PubMed: 18142378.
- Jaaro H, Beck G, Conticello SG, Fainzilber M (2001) Evolving better brains: a need for neurotrophins? *Trends Neurosci* 24: 79–85. doi:10.1016/S0166-2236(00)01690-8. PubMed: 11164937.
- Huang EJ, Reichardt LF (2001) Neurotrophins: roles in neuronal development and function. *Annu Rev Neurosci* 24: 677–736. doi:10.1146/annurev.neuro.24.1.677. PubMed: 11520916.
- Korsching S (1993) The neurotrophic factor concept: a reexamination. *J Neurosci* 13: 2739–2748. PubMed: 8331370.
- Lewin GR, Barde YA (1996) Physiology of the neurotrophins. *Annu Rev Neurosci* 19: 289–317. doi:10.1146/annurev.ne.19.030196.001445. PubMed: 8833445.
- Lindsay RM, Wiegand SJ, Altar CA, DiStefano PS (1994) Neurotrophic factors: from molecule to man. *Trends Neurosci* 17: 182–190. doi:10.1016/0166-2236(94)90099-X. PubMed: 7520198.
- Barbacid M (1994) The Trk family of neurotrophin receptors. *J Neurobiol* 25: 1386–1403. doi:10.1002/neu.480251107. PubMed: 7852993.
- Meakin SO, Shooter EM (1992) The nerve growth factor family of receptors. *Trends Neurosci* 15: 323–331. doi:10.1016/0166-2236(92)90047-C. PubMed: 1382329.
- Kaplan DR, Hempstead BL, Martin-Zanca D, Chao M V, Parada LF (1991) The trk proto-oncogene product: a signal transducing receptor for nerve growth factor. *Science* 252: 554–558.
- Klein R, Nanduri V, Jing SA, Lamballe F, Tapley P et al. (1991) The trkB tyrosine protein kinase is a receptor for brain-derived neurotrophic factor and neurotrophin-3. *Cell* 66: 395–403. doi:10.1016/0092-8674(91)90628-C. PubMed: 1649702.
- Hallböök F, Lundin LG, Kullander K (1998) *Lampetra fluviatilis* neurotrophin homolog, descendant of a neurotrophin ancestor,

- discloses the early molecular evolution of neurotrophins in the vertebrate subphylum. *J Neurosci* 18: 8700–8711. PubMed: 9786977.
41. Koyama J, Inoue S, Ikeda K, Hayashi K (1992) Purification and amino-acid sequence of a nerve growth factor from the venom of *Vipera russelli russelli*. *Biochim Biophys Acta* 1160: 287–292. doi: 10.1016/0167-4838(92)90090-Z. PubMed: 1477101.
 42. Sanz L, Gibbs HL, Mackessy SP, Calvete JJ (2006) Venom proteomes of closely related *Sistrurus rattlesnakes* with divergent diets. *J Proteome Res* 5: 2098–2112. doi:10.1021/pr0602500. PubMed: 16944921.
 43. Siigur J, Arumäe U, Neuman T, Siigur E, Saarma M (1987) Monoclonal antibody immunoaffinity chromatography of the nerve growth factor from snake venoms. *Comp Biochem Physiol B* 87: 329–334. PubMed: 3621902.
 44. Earl ST, Birrell GW, Wallis TP, St Pierre LD, Masci PP et al. (2006) Post-translational modification accounts for the presence of varied forms of nerve growth factor in Australian elapid snake venoms. *Proteomics* 6: 6554–6565. doi:10.1002/pmic.200600263. PubMed: 17109379.
 45. Morrison JJ, Pearn JH, Coulter AR (1982) The mass of venom injected by two elapidae: the taipan (*Oxyuranus scutellatus*) and the Australian tiger snake (*Notechis scutatus*). *Toxicon* 20: 739–745. doi: 10.1016/0041-0101(82)90121-0. PubMed: 6753241.
 46. Morrison JJ, Tesseraux I, Pearn JH, Harris J, Masci PP (1984) Venom of the Australian rough-scaled snake, *Tropidechis carinatus*: lethal potency and electrophysiological actions. *Toxicon* 22: 759–765. doi: 10.1016/0041-0101(84)90158-2. PubMed: 6098054.
 47. Ching AT, Paes Leme AF, Zelanis A, Rocha MM, Furtado Mde F et al. (2012) Venomics profiling of *Thamnodynastes strigatus* unveils matrix metalloproteinases and other novel proteins recruited to the toxin arsenal of rear-fanged snakes. *J Proteome Res* 11: 1152–1162. doi: 10.1021/pr200876c. PubMed: 22168127.
 48. Ching AT, Rocha MM, Paes Leme AF, Pimenta DC, de Fatima DFM et al. (2006) Some aspects of the venom proteome of the Colubridae snake *Philodryas* *offersii* revealed from a Duvernoy's (venom) gland transcriptome. *FEBS Lett* 580: 4417–4422. doi:10.1016/j.febslet.2006.07.010. PubMed: 16857193.
 49. Fry BG, Casewell NR, Wüster W, Vidal N, Young B et al. (2012) The structural and functional diversification of the *Toxicofera* reptile venom system. *Toxicon* 60: 434–448. doi:10.1016/j.toxicon.2012.02.013. PubMed: 22446061.
 50. Fry BG, Scheib H, de LMJ de Al, Silva DA, Casewell NR (2012) Novel transcripts in the maxillary venom glands of advanced snakes. *Toxicon* 59: 696–708. doi:10.1016/j.toxicon.2012.03.005. PubMed: 22465490.
 51. Fry BG, Lumsden NG, Wüster W, Wickramaratna JC, Hodgson WC et al. (2003) Isolation of a neurotoxin (alpha-colubritoxin) from a nonvenomous colubrid: evidence for early origin of venom in snakes. *J Mol Evol* 57: 446–452. doi:10.1007/s00239-003-2497-3. PubMed: 14708577.
 52. Sunagar K, Johnson WE, O'Brien SJ, Vasconcelos V, Antunes A (2012) Evolution of CRISPs associated with toxicofera-reptilian venom and mammalian reproduction. *Mol Biol Evol* 29: 1807–1822. doi: 10.1093/molbev/mss058. PubMed: 22319140.
 53. Snider WD (1994) Functions of the neurotrophins during nervous system development: what the knockouts are teaching us. *Cell* 77: 627–638. doi:10.1016/0092-8674(94)90048-5. PubMed: 8205613.
 54. Nisani Z, Dunbar SG, Hayes WK (2007) Cost of venom regeneration in *Parabuthus transvaalicus* (Arachnida: Buthidae). *Comp Biochem Physiol A Mol Integr Physiol* 147: 509–513. doi:10.1016/j.cbpa.2007.01.027. PubMed: 17344080.
 55. Nisani Z, Boskovic DS, Dunbar SG, Kelln W, Hayes WK (2012) Investigating the chemical profile of regenerated scorpion (*Parabuthus transvaalicus*) venom in relation to metabolic cost and toxicity. *Toxicon* 60: 315–323. doi:10.1016/j.toxicon.2012.04.343. PubMed: 22564718.
 56. Morgenstern D, King GF (2013) The venom optimization hypothesis revisited. *Toxicon* 63: 120–128. doi:10.1016/j.toxicon.2012.11.022. PubMed: 23266311.
 57. Brust A, Sunagar K, Undheim EAB, Vetter I, Yang DC et al. (2013) Differential evolution and neofunctionalization of snake venom metalloprotease domains. *Mol Cell Proteomics* 12: 651–663. doi: 10.1074/mcp.M112.023135. PubMed: 23242553.
 58. Kini RM, Chan YM (1999) Accelerated evolution and molecular surface of venom phospholipase A2 enzymes. *J Mol Evol* 48: 125–132. doi: 10.1007/PL00006450. PubMed: 9929380.
 59. Ruder T, Sunagar K, Undheim EAB, Ali SA, Wai T-C et al. (2013) Molecular phylogeny and evolution of the proteins encoded by coleoid (cuttlefish, octopus, and squid) posterior venom glands. *J Mol Evol* 76: 192–204. doi:10.1007/s00239-013-9552-5. PubMed: 23456102.
 60. Kini RM (2003) Excitement ahead: structure, function and mechanism of snake venom phospholipase A2 enzymes. *Toxicon* 42: 827–840. doi: 10.1016/j.toxicon.2003.11.002. PubMed: 15019485.
 61. Kini RM (2005) Structure-function relationships and mechanism of anticoagulant phospholipase A2 enzymes from snake venoms. *Toxicon* 45: 1147–1161. doi:10.1016/j.toxicon.2005.02.018. PubMed: 15922780.
 62. Low DHW, Sunagar K, Undheim EAB, Ali SA, Alagon AC et al. (2013) Dracula's children: Molecular evolution of vampire bat venom. *J Proteomics* 89: 95–111. doi:10.1016/j.jprot.2013.05.034. PubMed: 23748026.
 63. Wijeyewickrema LC, Gardiner EE, Gladigau EL, Berndt MC, Andrews RK (2010) Nerve growth factor inhibits metalloproteinase-disintegrins and blocks ectodomain shedding of platelet glycoprotein. *Vi - J Biol Chem* 285: 11793–11799. doi:10.1074/jbc.M110.100479.
 64. Cohen S (1959) Purification and metabolic effects of a nerve growth-promoting protein from snake venom. *J Biol Chem* 234: 1129–1137. PubMed: 13654333.
 65. Wijeyewickrema LC, Gardiner EE, Shen Y, Berndt MC, Andrews RK (2007) Fractionation of snake venom metalloproteinases by metal ion affinity: a purified cobra metalloproteinase, Nk, from *Naja kaouthia* binds Ni²⁺-agarose. *Toxicon* 50: 1064–1072. doi:10.1016/j.toxicon.2007.07.006. PubMed: 17822731.
 66. Frade JM, Barde YA (1999) Genetic evidence for cell death mediated by nerve growth factor and the neurotrophin receptor p75 in the developing mouse retina and spinal cord. *Development* 126: 683–690. PubMed: 9895316.
 67. Frade JM, Barde YA (1998) Microglia-derived nerve growth factor causes cell death in the developing retina. *Neuron* 20: 35–41. doi: 10.1016/S0896-6273(00)80432-8. PubMed: 9459440.
 68. Frade JM, Rodríguez-Tébar A, Barde YA (1996) Induction of cell death by endogenous nerve growth factor through its p75 receptor. *Nature* 383: 166–168. doi:10.1038/383166a0. PubMed: 8774880.
 69. Girish KS, Jagadeesha DK, Rajeev KB, Kemparaju K (2002) Snake venom hyaluronidase: an evidence for isoforms and extracellular matrix degradation. *Mol Cell Biochem* 240: 105–110. doi:10.1023/A:1020651607164. PubMed: 12487377.
 70. Lindholm D, Heumann R, Meyer M, Thoenen H (1987) Interleukin-1 regulates synthesis of nerve growth factor in non-neuronal cells of rat sciatic nerve. *Nature* 330: 658–659. doi:10.1038/330658a0. PubMed: 3317065.
 71. Woolf CJ (1996) Phenotypic modification of primary sensory neurons: the role of nerve growth factor in the production of persistent pain. *Philos Trans R Soc L B Biol Sci* 351: 441–448. doi:10.1098/rstb.1996.0040.
 72. Theoharides TC (1990) Mast cells: the immune gate to the brain. *Life Sci* 46: 607–617. doi:10.1016/0024-3205(90)90129-F. PubMed: 2407920.
 73. Payne V, Kam PC (2004) Mast cell tryptase: a review of its physiology and clinical significance. *Anaesthesia* 59: 695–703. doi:10.1111/j.1365-2044.2004.03757.x. PubMed: 15200544.
 74. Katzir I, Shani J, Goshen G, Sela J, Ninary E et al. (2003) Characterization of nerve growth factors (NGFs) from snake venoms by use of a novel, quantitative bioassay utilizing pheochromocytoma (PC12) cells overexpressing human trkA receptors. *Toxicon* 42: 481–490. doi:10.1016/S0041-0101(03)00225-3. PubMed: 14529729.
 75. Pearce FL, Banks BE, Banthorpe DV, Berry AR, Davies HS et al. (1972) The isolation and characterization of nerve-growth factor from the venom of *Vipera russelli*. *Eur J Biochem* 29: 417–425. doi: 10.1111/j.1432-1033.1972.tb02004.x. PubMed: 4563340.
 76. Pearce FL, Thompson HL (1986) Some characteristics of histamine secretion from rat peritoneal mast cells stimulated with nerve growth factor. *J Physiol* 372: 379–393. PubMed: 2425086.
 77. Marshall RD (1972) Glycoproteins. *Annu Rev Biochem* 41: 673–702. doi:10.1146/annurev.bi.41.070172.003325. PubMed: 4563441.
 78. Gavel Y, von Heijne G (1990) Sequence differences between glycosylated and non-glycosylated Asn-X-Thr/Ser acceptor sites: implications for protein engineering. *Protein Eng* 3: 433–442. doi: 10.1093/protein/3.5.433. PubMed: 2349213.
 79. Kornfeld R, Kornfeld S (1985) Assembly of asparagine-linked oligosaccharides. *Annu Rev Biochem* 54: 631–664. doi:10.1146/annurev.bi.54.070185.003215. PubMed: 3896128.
 80. Koh DC, Armugam A, Jeyaseelan K (2004) Sputa nerve growth factor forms a preferable substitute to mouse 7S-beta nerve growth factor. *Biochem J* 383: 149–158. doi:10.1042/BJ20040569. PubMed: 15225125.
 81. Casewell NR, Huttley GA, Wüster W, Wüster W (2012) Dynamic evolution of venom proteins in squamate reptiles. *Nat Commun* 3: 1066. doi:10.1038/ncomms2065. PubMed: 22990862.

82. Townsend TM, Mulcahy DG, Noonan BP, Sites JW Jr., Kuczyński CA et al. (2011) Phylogeny of iguanian lizards inferred from 29 nuclear loci, and a comparison of concatenated and species-tree approaches for an ancient, rapid radiation. *Mol Phylogenet Evol* 61: 363–380. doi:10.1016/j.jmpev.2011.07.008. PubMed: 21787873.
83. Wiens JJ, Hutter CR, Mulcahy DG, Noonan BP, Townsend TM et al. (2012) Resolving the phylogeny of lizards and snakes (Squamata) with extensive sampling of genes and species. *Biol Lett* 8: 1043–1046. doi:10.1098/rsbl.2012.0703. PubMed: 22993238.
84. Wiens JJ, Kuczyński CA, Smith SA, Mulcahy DG, Sites JW Jr. et al. (2008) Branch lengths, support, and congruence: testing the phylogenomic approach with 20 nuclear loci in snakes. *Syst Biol* 57: 420–431. doi:10.1080/10635150802166053. PubMed: 18570036.
85. Wiens JJ, Kuczyński CA, Townsend T, Reeder TW, Mulcahy DG et al. (2010) Combining phylogenomics and fossils in higher-level squamate reptile phylogeny: molecular data change the placement of fossil taxa. *Syst Biol* 59: 674–688. doi:10.1093/sysbio/syq048. PubMed: 20930035.
86. Casewell NR, Wagstaff SC, Harrison RA, Wüster W (2011) Gene tree parsimony of multilocus snake venom protein families reveals species tree conflict as a result of multiple parallel gene loss. *Mol Biol Evol* 28: 1157–1172. doi:10.1093/molbev/msq302. PubMed: 21062752.
87. Edgar RC (2004) MUSCLE: Multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res* 32: 1792–1797. doi:10.1093/nar/gkh340. PubMed: 15034147.
88. Ronquist F, Teslenko M, van der Mark P, Ayres DL, Darling A et al. (2012) MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. *Syst Biol* 61: 539–542. doi:10.1093/sysbio/sys029. PubMed: 22357727.
89. Guindon S, Dufayard JF, Lefort V, Anisimova M, Hordijk W, et al. (2010) New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Syst Biol* 59: 307–321. doi:10.1093/sysbio/syq010.
90. Posada D, Crandall KA (2002) The effect of recombination on the accuracy of phylogeny estimation. *J Mol Evol* 54: 396–402. doi:10.1007/s00239-001-0034-9. PubMed: 11847565.
91. Delpont W, Poon AF, Frost SD, Kosakovsky Pond SL (2010) Datamonkey 2010: a suite of phylogenetic analysis tools for evolutionary biology. *Bioinformatics* 26: 2455–2457. doi:10.1093/bioinformatics/btq429. PubMed: 20671151.
92. Kosakovsky Pond SL, Posada D, Gravenor MB, Woelk CH, Frost SDW et al. (2006) Automated phylogenetic detection of recombination using a genetic algorithm. *Mol Biol Evol* 23: 1891–1901. doi:10.1093/molbev/msl051. PubMed: 16818476.
93. Goldman N, Yang Z (1994) A codon-based model of nucleotide substitution for protein-coding DNA sequences. *Mol Biol Evol* 11: 725–736. PubMed: 7968486.
94. Yang Z (1998) Likelihood ratio tests for detecting positive selection and application to primate lysozyme evolution. *Mol Biol Evol* 15: 568–573. doi:10.1093/oxfordjournals.molbev.a025957. PubMed: 9580986.
95. Yang Z (2007) PAML 4: Phylogenetic analysis by maximum likelihood. *Mol Biol Evol* 24: 1586–1591. doi:10.1093/molbev/msm088. PubMed: 17483113.
96. Yang Z, Nielsen R (2002) Codon-substitution models for detecting molecular adaptation at individual sites along specific lineages. *Mol Biol Evol* 19: 908–917. doi:10.1093/oxfordjournals.molbev.a004148. PubMed: 12032247.
97. Zhang J, Nielsen R, Yang Z (2005) Evaluation of an improved branch-site likelihood method for detecting positive selection at the molecular level. *Mol Biol Evol* 22: 2472–2479. doi:10.1093/molbev/msi237. PubMed: 16107592.
98. Nielsen R, Yang Z (1998) Likelihood models for detecting positively selected amino acid sites and applications to the HIV-1 envelope gene. *Genetics* 148: 929–936. PubMed: 9539414.
99. Kosakovsky Pond SL, Frost SDW, Pond SLK (2005) Not so different after all: a comparison of methods for detecting amino acid sites under selection. *Mol Biol Evol* 22: 1208–1222. doi:10.1093/molbev/msi105. PubMed: 15703242.
100. Murrell B, Moola S, Mabona A, Weighill T, Sheward D et al. (2013) FUBAR: a fast, unconstrained bayesian approximation for inferring selection. *Mol Biol Evol* 30: 1196–1205. doi:10.1093/molbev/mst030. PubMed: 23420840.
101. Pond SLK, Frost SDW, Muse SV (2005) HyPhy: hypothesis testing using phylogenies. *Bioinformatics* 21: 676–679. doi:10.1093/bioinformatics/bti079. PubMed: 15509596.
102. Murrell B, Wertheim JO, Moola S, Weighill T, Scheffler K et al. (2012) Detecting individual sites subject to episodic diversifying selection. *PLoS Genet* 8: e1002764. doi:10.1371/journal.pgen.1002764. PubMed: 22807683.
103. Woolley S, Johnson J, Smith MJ, Crandall KA, McClellan DA (2003) TreeSAAP: selection on amino acid properties using phylogenetic trees. *Bioinformatics* 19: 671–672. doi:10.1093/bioinformatics/btg043. PubMed: 12651734.
104. Bielawski JP, Yang Z (2004) A maximum likelihood method for detecting functional divergence at individual codon sites, with application to gene family evolution. *J Mol Evol* 59: 121–132. doi:10.1007/s00239-004-2597-8. PubMed: 15383915.
105. Pond SL, Frost SD (2005) Datamonkey: rapid detection of selective pressure on individual sites of codon alignments. *Bioinformatics* 21: 2531–2533. doi:10.1093/bioinformatics/bti320. PubMed: 15713735.
106. Pond SLK, Scheffler K, Gravenor MB, Poon AFY, Frost SDW (2010) Evolutionary fingerprinting of genes. *Mol Biol Evol* 27: 520–536. doi:10.1093/molbev/msp260. PubMed: 19864470.
107. Kelley LA, Sternberg MJE (2009) Protein structure prediction on the Web: a case study using the Phyre server. *Nat Protoc* 4: 363–371. doi:10.1038/nprot.2009.2. PubMed: 19247286.
108. DeLano WL (2002) The PyMOL Molecular Graphics System. DeLano. Sci San Carlos, CA.
109. Armon A, Graur D, Ben-Tal N (2001) ConSurf: an algorithmic tool for the identification of functional regions in proteins by surface mapping of phylogenetic information. *J Mol Biol* 307: 447–463. doi:10.1006/jmbi.2000.4474. PubMed: 11243830.
110. Fraczekiewicz R, Braun W (1998) Exact and efficient analytical calculation of the accessible surface areas and their gradients for macromolecules. *J Comput Chem* 19: 319–333.